

# opendata to ECHO10 mapping

See Schema documentation at <http://project-open-data.github.io/schema/> and mapping to ISO19115 and other standard schemas at <http://project-open-data.github.io/metadata-resources/#common-core-required-fields-equivalents>

## “Common Core” Required Fields

| Field        | Definition  | ECHO 10 Mapping   |
|--------------|---|---|
| title        | Human-readable name of the asset. Should be in plain English and include sufficient detail to facilitate search and discovery.  | DataSetId   |
| description  | Human-readable description (e.g., an abstract) with sufficient detail to enable a user to quickly understand whether the asset is of interest.  | <Description>   |
| keyword      | Tags (or keywords) help users discover your dataset; please include terms that would be used by technical and non-technical users.  | For each ScienceKeyword:<br>category >> topic >> term<br><br><i>This format was derived from Reverb. Note that we don't include VariableLevel1Keyword or Spatial/Temporal keywords at this time as they are rarely used. If there are no ScienceKeywords, the opendata record will be invalid, though it is not clear how common it is to have no Science Keywords.</i>   |
| modified     | Most recent date on which the dataset was changed, updated or modified.   | LastUpdate  |
| publisher    | The publishing entity.  | ArchiveCenter   |
| contactPoint | Contact person's name for the asset.  | One of: <ul style="list-style-type: none"><li>• ContactPerson/FirstName[1] ContactPerson/LastName[1] (Contact/OrganizationName[1])</li><li>• ContactPerson/FirstName[1] ContactPerson/LastName[1]</li><li>• Contact/OrganizationName[1]</li><li>• Contact/Role[1]</li></ul> In order of priority depending on available metadata.<br><br><i>Note: will always use the first returned value which may or may not be the most appropriate contact</i>                   |
| mbox         | Contact person's email address.   | OrganizationEmails/Email[1]<br><br><i>Note: will always use the first returned value which may or may not be the most appropriate email.</i>  |
| identifier   | A unique identifier for the dataset or API as maintained within an Agency catalog or database.  | echo_collection_id<br><br><i>Note: In the official ISO mapping, this maps to //gmd:fileIdentifier/gco:CharacterString which in the ECHO &lt;-&gt; ISO mapping is gov.nasa.echo:/Collection/DataSetId[1]. So, By that logic, we should use DatasetID here, but that is not unique (in OPS it SHOULD be, but we dont enforce uniqueness of this field). For the moment, will use echo_collection_id for echo10 collections and (indirectly) datasetID for ISO 19115</i> |
| accessLevel  | The degree to which this dataset <b>could</b> be made publicly-available, <i>regardless of whether it has been made available</i> . Choices: public (Data asset is or could be made publicly available to all without restrictions), restricted public (Data asset is available under certain use restrictions), or non-public (Data asset is not available to members of the public) | 'public' for all collections except collections whose short names contain 'AST' which will be 'restricted public'   |

## “Common Core” Required-if-Applicable Fields

| Field              | Definition   | ECHO 10 Mapping  |
|--------------------|--|--|
| bureauCode         | Federal agencies, combined agency and bureau code from <a href="#">OMB Circular A-11, Appendix C</a> in the format of 015:11.  | '026:00' (NASA)<br><br><i>NOTE: in some cases a different bureau code(e.g. USGS) may be more appropriate. When these are identified, we will modify the mapping.</i>   |
| programCode        | Federal agencies, list the primary program related to this data asset, from the <a href="#">Federal Program Inventory</a> . Use the format of 015:001  | '026:001' (NASA : Earth Science Research)<br><br><i>NOTE: in some cases a different program code may be more appropriate. When these are identified, we will modify the mapping.</i>   |
| accessLevelComment | An explanation for the selected “accessLevel” including instructions for how to access a restricted file, if applicable, or explanation for why a “non-public” or “restricted public” data asset is not “public,” if applicable. Text, 255 characters. | RestrictionComment   |
| accessURL          | URL providing direct access to the downloadable distribution of a dataset.   | First OnlineAccessURL if it exists, otherwise, none.   |
| webService         | Endpoint of web service to access dataset.   | nil  |
| format             | The file format or API type of the distribution.   | The MimeType of the first OnlineAccessURL element, if any. nil otherwise.  |
| license            | The license with which the dataset or API is published. See <a href="#">Open Licenses</a> for more information.  | nil  |
| spatial            | The range of spatial applicability of a dataset. Could include a spatial region like a bounding box or a named place.  | One of: <ul style="list-style-type: none"><li>• "west, south, east, north" from bounding rectangle</li><li>• "x,y" from point</li></ul><br><i>NOTE: Currently this picks the first box or point, and multiple boxes or points will be ignored. Also, it does not support polygons and lines. This will work for most collections, but an NCR (11014593) has been written to add support for these other cases.</i> |
| temporal           | The range of temporal applicability of a dataset (i.e., a start and end date of applicability for the data).   | One of: <ul style="list-style-type: none"><li>• //Temporal/SingleDateTime</li><li>• //Temporal/RangeDateTime/BeginningDateTime '/ //Temporal/RangeDateTime/EndingDateTime</li></ul><br>Depending on temporal metadata of collection.<br><br><i>NOTE: Theoretically need to handle PeriodicDateTime as well, though We have no instances of PeriodicDateTime data now. This has been documented in NCR 11014594</i> |

## Expanded Fields

| Field | Definition                             | ECHO 10 Mapping  |
|-------|--|--|
| theme | Main thematic category of the dataset. | //Campaign/ShortName<br><br>If multiple campaigns, join with a ',' |

|                        |  |   |
|------------------------|--|---|
| dataDictionary         | URL to the data dictionary for the dataset or API. Note that documentation other than a data dictionary can be referenced using Related Documents as shown in the expanded fields. | nil   |
| dataQuality            | Whether the dataset meets the agency's Information Quality Guidelines (true/false).  | nil   |
| distribution           | Holds multiple download URLs for datasets composed of multiple files and/or file types   | <p>Array containing URL to search this collection in reverb (with MIME type text/html), plus any OnlineAccessURL (with the Mime type form theMimeType field, or 'application/octet-stream' if not set).</p> <p><i>NOTE: OnlineAccessURL is generally not populated for ECS data at least. In other cases (e.g. ASF&lt; it appears to be populated with a URL to a generic provider portal to get the data, not the data itself.</i></p>   |
| accrualPeriodicity     | Frequency with which dataset is published.   | <p>Map (upcased) MaintenanceAndUpdateFrequency to <a href="http://www.ukoln.ac.uk/metadata/dcmi/collection-DCCDAccrualPeriodicity/">http://www.ukoln.ac.uk/metadata/dcmi/collection-DCCDAccrualPeriodicity/</a> as follows</p> <p>'CONTINUALLY' =&gt; 'Continuously updated',<br/>     'DAILY' =&gt; 'Daily',<br/>     'ONCE A WEEK' =&gt; 'Weekly',<br/>     'WEEKLY' =&gt; 'Weekly',<br/>     'MONTHLY' =&gt; 'Monthly',<br/>     'IRREGULAR' =&gt; 'Completely irregular',<br/>     'AS NEEDED' =&gt; 'Completely irregular',<br/>     'NONE PLANNED' =&gt; nil,<br/>     'NONE' =&gt; nil</p> <p>Unrecognized values will map to nil.</p> |
| landingPage            | Alternative landing page used to redirect user to a contextual, Agency-hosted "homepage" for the Dataset or API when selecting this resource from the Data.gov user interface.     | <p>URL to search this collection in reverb.</p> <p><i>NOTE: The detailed API description says</i></p> <p style="margin-left: 20px;">"This allows agencies to better specify what a visitor receives after selecting one of the agency's datasets on Data.gov or in third-party mashups."</p>  |
| language               | The language of the dataset.   | 'en-US'   |
| PrimaryITInvestmentUII | For linking a dataset with an IT Unique Investment Identifier (UII)  | nil   |
| references             | Related documents such as technical information about a dataset, developer documentation, etc.   | list of all OnlineResource/URL values   |
| issued                 | Date of formal issuance.   | InsertTime  |
| systemOfRecords        | If the system is designated as a system of records under the Privacy Act of 1974, provide the URL to the System of Records Notice related to this dataset.                         | nil   |

## Simple conversion stylesheet

Note that the catalog-rest code does not use this stylesheet. This is only meant as a testing tool to see what values catalog-rest should be pulling based on the above mapping.

Run the following on the command line after saving the stylesheet to a file to apply it to a metadata file:

```
xsltproc <XSLT File> <Metadata File> | xmllint --format -
```

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0" xmlns="http://gcmd.gsfc.nasa.gov/Aboutus
/xml/dif" xmlns:dif="http://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" x
si:schemaLocation="http://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/http://gcmd.gsfc.nasa.gov/Aboutus/xml/dif/dif_v9.8.4.xsd">

<xsl:template match='/'>
<opendata>
<title>
<xsl:value-of select="//DataSetId"/>
</title>
<description>
<xsl:value-of select="//Description"/>
(Suggested Usage: <xsl:value-of select="//SuggestedUsage"/>)
</description>
<keywords>
<xsl:for-each select="//ScienceKeywords/ScienceKeyword">
<keyword><xsl:value-of select="CategoryKeyword"/> >> <xsl:value-of select="TopicKeyword"/> >> <xsl:value-of select="Ter
mKeyword"/></keyword>
</xsl:for-each>
</keywords>
<modified>
<xsl:value-of select="//LastUpdate"/>
</modified>
<publisher>
<xsl:value-of select="//ArchiveCenter"/>
</publisher>
<contactPoint>
<xsl:value-of select="//ContactPerson/FirstName"/> <xsl:value-of select="//ContactPerson/LastName"/> (<xsl:value-of select
="//ContactPerson/OrganizationName"/>)
<xsl:value-of select="//Contact/Role"/>
</contactPoint>
<mbox>
<xsl:value-of select="//OrganizationEmails/Email"/>
</mbox>
<accessLevelComment>
<xsl:value-of select="//RestrictionComment"/>
</accessLevelComment>
<spatial>
<xsl:value-of select="//Geometry/BoundingRectangle/WestBoundingCoordinate"/>, <xsl:value-of select="//Geometry/Boundi
ngRectangle/SouthBoundingCoordinate"/>, <xsl:value-of select="//Geometry/BoundingRectangle/EastBoundingCoordinate"/
>, <xsl:value-of select="//Geometry/BoundingRectangle/NorthBoundingCoordinate"/>
</spatial>
<temporal>
<xsl:choose>
<xsl:when test="count(/Temporal/SingleDateTime) > 0">
<xsl:value-of select="//Temporal/SingleDateTime"/>
</xsl:when>
<xsl:when test="count(/Temporal/RangeDateTime) > 0">
<xsl:value-of select="//Temporal/RangeDateTime/BeginningDateTime"/> / <xsl:value-of select="//Temporal/RangeDateTime/
EndingDateTime"/>
</xsl:when>
</xsl:choose>
</temporal>
<accrualPeriodicity>
<xsl:variable name="period" select="//MaintenanceAndUpdateFrequency"/>
<xsl:variable name="upperCase">ABCDEFGHIJKLMNPQRSTUVWXYZ</xsl:variable>
<xsl:variable name="lowerCase">abcdefghijklmnopqrstuvwxyz</xsl:variable>
<xsl:choose>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'CONTINUALLY'">
Continuously updated
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'DAILY'">
Daily
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'ONCE A WEEK'">
Weekly
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'WEEKLY'">
```

```
Weekly
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'MONTHLY'>
Monthly
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'IRREGULAR'">
Completely irregular
</xsl:when>
<xsl:when test="translate($period,$lowerCase,$upperCase) = 'AS NEEDED'">
Completely irregular
</xsl:when>
</xsl:choose>
</accrualPeriodicity>
<references>
<xsl:for-each select="//OnlineResource/URL">
<reference><xsl:value-of select='.'/></reference>
</xsl:for-each>
</references>
<issued>
<xsl:value-of select="//InsertTime"/>
</issued>
</opendata>
</xsl:template>
</xsl:stylesheet>
```